DOCUMENT RESUME

ED 269 084 JC 860 229

AUTHOR Grosset, Jane M.; Hawk, Thomas R.
TITLE Development of a Socio-Economic-Status Index Using United States Census Data.
PUB DATE Apr 86
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Census Figures; *Economic Factors; Factor Analysis; High Risk Students; Income; *Predictor Variables; *Research Methodology; *Social Indicators; Socioeconomic Background; *Socioeconomic Status; Two Year Colleges; *Two Year College Students; Urban Education; Withdrawal (Education)

ABSTRACT

The study reported here employed a quasi-factorial ecological approach to explore the possibility of using economic and social indicators available from the 1980 census to construct a socio-economic status (SES) index. The study hypothesized that if an appropriate factor analysis model could be identified, factor score equations could be used to construct an SES index for each of the 49 zip codes in a large eastern city. In turn, students attending a public two-year college in the city could be assigned an SES measure on the basis of their residential zip code. Using census data, seven economic and social indicators (i.e., mean per-capita income, median household income, percentage of persons below the poverty level, percentage of high-school graduates and college graduates, percentage of individuals on unemployment, and percentage of individuals on work disability) were factor analyzed. Indices representing dimensions of economic deprivation and educational attainment were used to categorize zip codes as either high or low SES. Correlations of the SES indices with high-risk student behaviors, academic performance, and attrition were statistically significant in most analyses, albeit they offered only modest contributions to the explanation of overall variation. Institutional efforts to use the SES variables in counseling and recruitment have been extremely cautious, in recognition of the fact that SES data can only be used in the context of a holistic assessment. (RO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. GROSSET

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☒ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED269084

DEVELOPMENT OF A SOCIO-ECONOMIC-STATUS INDEX USING UNITED STATES CENSUS DATA

Jane M. Grosset

Thomas R. Hawk

Community College of Philadelphia
1700 Spring Garden Street
Philadelphia, PA 19130

IC 860 229

Presented at the 70th Annual Meeting of the

American Educational Research Association

San Francisco, California

April 16-20, 1986

The ability to identify "high risk" college students and intervene with an appropriate combination of recommended or mandated course selections and support services is usually considered to be essential to two-year college effectiveness. Over ninety percent of public two-year colleges now test entering students (Woods, 1985). However, placement and admissions tests scores alone are not sufficient to provide a detailed understanding of an entering student's need for instructional support. Institutional research by the authors, as well as published research findings, have suggested that accurate counseling and program placement can best occur in the context of a holistic assessment which considers, in addition to test results, data on a student's prior education experiences and achievements, their educational aspirations, and socio-economic background.

Tinto (1975), in an attempt to summarize existing literature on the factors associated with college dropout, suggested that retention is both a function of a student's academic and social integration into a college environment. Reliance on test results alone in placing students in courses and support services results in a failure to consider the other dimensions of a student's background which may impact upon the student's retention and academic success. While no one variable has emerged as consistently reliable, research to date has generally supported

the notion that students from a socio-economically and educationally disadvantaged background have more difficulty adopting to the expectations of higher education and therefore have a higher probability of poor academic performance and dropping out before completing educational goals. (See for example, Astin et al., 1972; Astin & Panos, 1969; Donovan, 1984; Gosman et al., 1983; Pantages & Creedon, 1978; Wilson, 1980.) This study assumed that these findings were sufficiently compelling to include a consideration of socio-economic status (SES) in a holistic assessment. This paper is a description of an approach to attempt to define a viable measure of a student's socio-economic background for use in a two-year public college in a large East-Coast city.

The challenge for two-year public colleges is to gather usable socio-economic data within the prevailing admissions and registration procedures. Information collection for two-year public-college students is problematic for several reasons. Many students defer application until just before the term commences, making orderly compilation of complete assessment data impossible. Enrollments during last-moment, walk-in registration sessions are common. Timely access by counselors and advisors to data which is collected is often difficult. Furthermore, a large percentage of two-year students are independent and are frequently reluctant to provide information regarding parents' educational and occupational backgrounds. Family income is an unreliable measure of the student's socio-economic status, since this is typically low for independent students irrespective of

their circumstances in their formative years. Still, the responsibility remains for the institution to gather sufficient clues from available information to make a reasoned judgement on the student's need for support during the critical first year of college enrollment.

Of the variables that could be collected for entering students in an expeditious fashion, zip code appeared to hold the most promise as a possible proxy for the student's SES background. The population density of the City served by the College has resulted in a large number of zip codes (49) within the College service area. This fact, coupled with the characteristic of the City to have relatively homogeneous neighborhoods, led to the hypothesis that zip codes could be used to help describe the students' SES backgrounds.

Research by Shevky and Bell (1955) supported the idea that zip codes can serve as a proxy for SES. They deduced there were three major dimensions of urban differentiation: social rank (economic status); urbanization (family status); and segregation (ethnic status). In their studies, Shevky and Bell used a factorial ecology approach which is characterized by the following: (1) interest in differentiation of residential areas within a metropolitan community; (2) extensive use of official census data as a primary source of variables; and (3) routine application of factor analysis procedures to explore census data (Jackson & Borgatta, 1981).

Methodology

In this study, a quasi-factorial ecological approach was used to explore the possibility of using economic and social indicators available from the 1980 Census of Population and Housing to construct an SES index. If an appropriate factor analysis model could be identified, factor score equations would be used to construct an SES index for each of the 49 zip codes. In turn, students would be assigned an SES measure based on their residential zip code.

Five categories of economic and social indicators, each with several associated measures, were available for all of the 49 zip codes. The indicators and measures appear in Table 1.

Table 1

Economic and Social Indicators
1980 Census of Population and Housing

Category	Measure
Population	Number of persons
	Number of families
Income	Per-capita income (mean)
	Family income (median)
	Household income (median)
Poverty	Persons below the poverty level (%)
	Families below the poverty level (%)
Education	High-school graduates (%)
	College graduates (%)
Civilian Labor Force	Unemployment level (%)
	Work disability level (%)
	Service industry (%)
	Retail-wholesale industry (%)
	Manufacturing (%)

With the exception of population, the measures associated with each of the categories fit conceptually within the SES construct. The original population measures were manipulated to compute a family-size variable for each zip code, which was used in the analysis instead of total persons and families.

Six of the 13 variables were eventually dropped from the analysis. Correlations of family income with per-capita income ($r = 0.93$) and families below the poverty level with persons below the poverty level ($r = 0.98$) were extremely high, indicating that the two sets of variables contained trivially redundant content (Jackson & Borgatta, 1981). As a consequence, two separate factor analyses were undertaken, each including one of the two sets of income and poverty variables. Since both sets of variables resulted in approximately the same results, per-capita income and persons below the poverty level were kept in ensuing analyses while family income and families below the poverty level were dropped out. Family size and the three industry-related variables dropped out of the model because of low sampling-adequacy measures and small squared multiple-correlation coefficients.

The remaining seven items, mean per-capita income, median household income, percentage of persons below the poverty level, percentage of high-school graduates and college graduates, percentage of individuals on unemployment and percentage of individuals on work disability,, were factor analyzed using the FACTOR subprogram in SPSS-X. An iterative procedure (Maximum Likelihood) was used to extract the initial values, applying an

eigenvalue criterion of 1.00 as the analysis termination point. Since the index scores constructed from the factors were subsequently to be used in discriminant analyses, the initial factor matrix was orthogonally rotated to its final solution using the varimax criterion.

Results

A two-factor solution is optimal based on the marginal eigenvalues criterion. The orthogonal solutions are presented in the following Table 2.

Table 2

Two Factor Solution: Loadings

Variable	Factor 1	Factor 2
% of persons below poverty	0.97756	-0.13891
Median household income	-0.86145	0.34837
% of unemployment	0.81309	-0.46203
% of disability	0.72568	-0.61875
% of college graduates	-0.09483	0.79996
% of high-school graduates	-0.51465	0.79284
Mean per-capita income	-0.46090	0.72750
Eigenvalue	4.737	1.170
Percent of variance	67.7	16.7
Cumulative percent	67.7	84.4

The four variables defining Factor 1, unemployment, disability, household income and poverty, relate to the level of economic deprivation in the zip-code area.

Factor 2, which includes per-capita income, high-school graduation and college graduation, can be thought of as providing a measure of educational attainments of zip-code residents.

Indices representing the two dimensions derived from the factor analysis were constructed for use in subsequent analyses. The method of constructing scales used only characteristic items for the factor rather than all items, irrespective of the size of their factor loadings. The Factor 1 score equation produced a scale ranging from -1.64 to 2.55. The scale for Factor 2 had a range between -1.44 and 3.15.

Efforts to Validate and Apply the SES Factor Scales

The scales and associated zip codes were shared with the City Planning Commission staff and College staff who were knowledgeable of City neighborhoods. Their assessment of the nature of Philadelphia neighborhoods paralleled that which was implied by the indices. Neighborhoods which they defined as homogeneous with regard to SES fell at the extremes of the indices. Heterogeneous neighborhoods had associated composite scores that fell in the mid-range. The results were consistent across the two indices with the same zip codes falling in the mid-range along both dimensions.

A major issue in using census data is the ten-year span between collections. A real concern was the reliability in 1985 of 1980 neighborhood information. In order to have some sense of the stability of the SES measures over time, 1970 and 1980 neighborhood profiles were developed and compared using Census economic and social indicators. This analysis helped to identify neighborhoods in transition whose SES values must be interpreted with caution.

Two strategies have been used in integrating the SES variables into institutional research. In most of the multivariate applications, the variables have been used as two separate scales in their original continuous form. For some applications, however, it has been advantageous to nominalize and collapse the two variables into one using the following scheme. Zip codes were dichotomized into either high (1) or low (0) SES categories based on their associated index scores. This approach produced a combination of 1's and 0's for each zip code, which were then collapsed into one trichotomized variable using the following criteria:

Factor 1	Factor 2	SES Level
1	0	high
0	1	middle
1	0	middle
0	0	low

While this approach yields a single SES variable that is less predictive than the two original indices, it has been extremely useful in providing a clear understanding of the interpretation of the variable to College staff.

The circumstances which resulted in the need for the SES proxy variables made it impossible to directly validate the reliability of the indices to measure actual students' backgrounds. Major effort was therefore focused on seeing how well the indices were related to student behavior while at the College. Several types of studies were undertaken:

1. Correlation of the SES indices with high-risk student behaviors, e.g., frequent course withdrawals,

inappropriate course selections and late registration.

2. Correlation of the SES indices with academic performance.
3. Correlation of SES indices with attrition before educational goals were accomplished.

In most analyses the SES variables emerged as statistically significant, albeit with contributions to overall variation explanation that were modest. Multivariate studies have shown the SES-proxy variables to make unique contributions to explaining student performance which cannot be attributed to other factors such as placement test scores, age or race.

Institutional efforts to utilize the SES variables have been appropriately cautious, with the knowledge that the SES variables provide limited additional information about the student that can only be used in the context of a holistic assessment. Counselors and advisors will use the SES variables as a cue to explore fully with the student needs for additional support services during their first semesters of enrollment at the College. The SES variables have also been used to begin to develop targeted recruitment efforts and pre-admission counseling approaches which recognize probable differences in potential students' understanding and expectations about college enrollment, and needs for assistance in order to make a meaningful enrollment decision.

BIBLIOGRAPHY

Astin, A.W. & Panos, R.J. The educational and vocational development of college students. Washington D.C.: American Council on Education, 1969.

Astin, H.S., Astin, A.W., Bisconti, A.S., & Frankel, H.H. Higher education and the disadvantaged student. Washington, D.C.: Human Service Press, 1972.

Donovan, R. Path analysis of a theoretical model of persistence in higher education among low-income black youth. Research in Higher Education, 1984, 21(1), 243-259.

Gosman, E.J., Dandridge, B.A., Nettles, M.T., & Thoeny, A.R. Predicting the student progression: the influence of race and other student and institutional characteristics on college student performance. Research in Higher Education, 1983, 18(2), 209-236.

Jackson, D. S. & Borgatta, E.F. Selecting a data analysis model for factorial ecology research. In D.S. Jackson and E.F. Borgatta (Eds.), Factor analysis and measurement in sociological research. Beverly Hills, CA.: Sage Press, 1981.

Pantages, T.J. & Creedon, C.F. Studies of college attrition: 1950-1975. Review of Educational Research, 1978, 48(1), 49-101.

Shevky, E. & Bell, W. Social area annalysis: illustrative application and computational procedures. Stanford, CA: Stanford University Press, 1955.

SPSS-X Users Guide. Chicago, IL: McGraw-Hill Book Co., 1983.

Tinto, V. Dropout from higher education: a theoretical synthesis of recent research. Review of Educational Research, 1975, 45(1), 89-125.

Wilson, K.M. The performance of minority students beyond the freshman year: testing a "late-bloomer" hypothesis in one state university setting. Research in Higher Education, 1980, 13(1), 23-46.

Woods, J.E. Status of testing practices at two-year postsecondary institutions. Washington, D.C.: American College Testing Program and the American Association of Community and Junior College, 1985.

ERIC Clearinghouse for Junior Colleges

JUN 13 1986